Cell, Vol. 74, 929–937, September 10, 1993, Copyright © 1993 by Cell Press

Prominent Role of Secondary Anchor Residues in Peptide Binding to HLA-A2.1 Molecules

Jörg Ruppert, John Sidney, Esteban Celis, Ralph T. Kubo, Howard M. Grey, and Alessandro Sette
Cytel Corporation
3525 John Hopkins Court
San Diego, California 92121

Summary

The functional determinants of histocompatibility leukocyte antigen (HLA)-A2.1–peptide interactions have been detailed by the use of quantitative molecular binding assays and a chemically synthesized library of naturally occurring epitopes. The importance of hydrophobic anchor residues in position 2 and the C-terminus was confirmed. These anchors are necessary, but not sufficient, for high affinity binding, as the predictions based solely on these anchors are only about 30% accurate. Prominent roles for several other positions (1, 3, and 7) were also demonstrated. The location of these residues within the peptides matches secondary A2.1 pockets previously demonstrated by X-ray crystallography. From a functional standpoint, similar dominant negative effects on binding were observed for charged residues in both nonamers and decamers, while positive effects differed between nonamers and decamers. An extended motif taking into account secondary anchors increased the predictability of A2.1-binding epitopes to a level of 70%, underscoring the practical usefulness of extended motifs.

Introduction

Presentation of antigenic peptides bound to major histocompatibility complex (MHC) class I molecules is a prerequisite for stimulation of cytotoxic T cell responses, a crucial defense mechanism against viral infections and tumors (Townsend and Bodmer, 1989; Bjorkman and Parham, 1990). Recent studies showed that the majority of peptides bound to class I molecules have a restricted size of 9 ± 1 amino acids and require free N- and C-terminal ends (Rötzschke et al., 1990; Elliot et al., 1991; Falk et al., 1991; Jardetzky et al., 1991; Schumacher et al., 1991). In addition to a specific size, different class I molecules appear to require a specific combination of usually two main anchor residues within their peptide ligands (Rötzschke and Falk, 1991). More specifically, in the case of the human allele histocompatibility leukocyte antigen (HLA)-A2.1, these anchors have been described as leucine (L) at position 2 and L or valine (V) at the C-terminal end (Falk et al., 1991). It is unclear, however, whether correct size and main anchor residues are, by themselves, sufficient for high affinity binding or whether other factors may also be important. Until now, the available data have been limited largely to the definition of which residues are most abundant at anchor positions, as determined by sequencing of pools of naturally processed peptides eluted from class I molecules (Falk et al., 1991). Systematic and quantitative studies on the structural requirements for peptide binding to class I are scarce (Chen and Parham, 1989; Corr et al., 1992). In particular, direct assessment of the effect of the presence of different side chains in the two main anchor positions, as well as an analysis of contributions of nonanchor residues to binding affinity, has not yet been reported.

At the level of the MHC molecules, a number of elegant X-ray crystallography studies have detailed the structure of both mouse and human class I molecules (Garret et al., 1989; Bjorkman and Parham, 1990; Saper et al., 1991; Fremont et al., 1992). Most interestingly, structures have been presented in which the binding groove was occupied by heterogenous, naturally processed material, as well as by class I molecules cocrystallized with single peptide entities (Jardetzky et al., 1991; Fremont et al., 1992; Matsumura et al., 1992). In the case of A2.1 molecules, in particular, detailed studies (Saper et al., 1991) have outlined the molecular structure of six different pockets in the peptide binding groove. The two main pockets (B and F) have been shown to engage the two main anchors located in position 2 and at the C-terminus of the peptide, respectively (Saper et al., 1991). By contrast, although no functional role has yet been determined for the remaining four pockets (the secondary pockets), the fact that they contain both polymorphic residues and extra electron density (Saper et al., 1991) suggests that they might play a significant role in A2.1–peptide interactions.

In this paper, large collections of synthetic peptides and quantitative direct HLA-A2.1 binding assays were used to investigate the role of amino acid residues other than the two main anchors (and secondary MHC pockets) in A2.1–peptide interactions.

Results

Single Amino Acid Substitutions of a Poly(A) 9-mer Peptide Carrying Anchor Residues at Positions 2 and 9

To investigate the structural requirements for the interaction of A2.1 molecules with their peptide ligands, single amino acid substitution analogs of a model A2.1-binding peptide were synthesized and tested for binding. The model peptide selected for this study was a poly(A) 9-mer peptide (ALAKAAAAV) containing the previously reported A2.1 motif L in position 2 (L_2) and V in position 9 (V_9) (Falk et al., 1991; Hunt et al., 1992). A lysine (K) residue was included in position 4 for the purpose of increasing the solubility of the peptide. The A2.1 binding capacity of this peptide and its various analogs were measured according to a recently described molecular binding assay that utilizes purified A2.1 molecules and radiolabeled synthetic peptides (A. S. et al., submitted). In this assay, we found a dissociation constant (K_D) value for the model A2.1 poly(A) peptide of 170 nM ± 72 (average of seven different experiments).

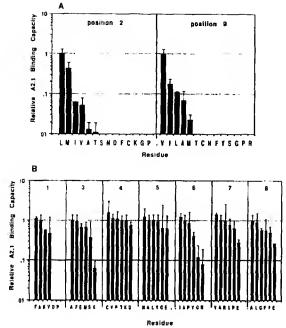

Figure 1. A2.1 Binding of Poly(A) Analogs

Single amino acid substitutions of the parental poly(A) 9-mer peptide ALAKAAAAV were tested for binding to purified HLA-A2.1 molecules. The indicated residues were introduced at the anchor positions 2 and 9 (A) or at nonanchor positions of the parental peptide (B). Binding is expressed relative to the parental unsubstituted peptide.

To obtain further information on the structural requirements of A2.1 binding, the binding capacity of a panel of single substitution analogs of this peptide was measured next. First, we sought to determine the degree of permissiveness of anchor positions 2 and 9. For this purpose, panels of 13 different analogs were synthesized for both anchor positions 2 and 9 (Figure 1A). In good agreement with the previously reported A2.1 motif, the peptides carrying L or methionine (M) in position 2 were the best binders. Marked decreases in binding capacity (10- to 100-fold) were apparent even with relatively conservative substitutions such as isoleucine (I), V, alanine (A), and threonine (T). More radical changes (i.e., residues aspartic acid [D], K, phenylalanine [F], cysteine [C], proline [P], glycine [G], asparagine [N], and serine [S]) completely abolished binding capacity. Similar results were obtained at position 9, where only conservative substitutions, such as L and I, bound within 10-fold of the unsubstituted model A2.1 peptide binder. Analogs carrying A or M substitutions also bound, but less strongly (10- to 100-fold decrease). Finally, all other substitutions tested (T, C, N, F, S, G, P, and arginine [R]) were associated with complete loss of A2.1 binding capacity. Thus, based on these data and in good agreement with previous studies (Falk et al., 1991; Hunt et al., 1992), a "canonical" A2.1 motif could be defined as L or M in position 2 and L, V, or I in position 9.

Next, to test whether positions other than 2 and 9 could also influence A2.1 binding capacity, five different single amino acid substituted analogs were synthesized at each of the nonanchor positions, 1, 3, 4, 5, 6, 7, and 8 (Figure 1B). When these analogs were tested for A2.1 binding, it was found that several of the substitutions also had significant effects on the A2.1 binding. At position 1, D and P substitutions virtually abolished A2.1 binding capacity, and at position 3, the positively charged K substitution led to a 15-fold reduction in A2.1 binding. Positions 4 and 5 were permissive, in the sense that none of the substitutions tested led to significant changes in binding capacity. At position 6, both a positively charged (R) and a G substitution substantially decreased binding. Finally, some detrimental effects (~4-fold) were also seen with a negatively charged substitution (glutamic acid [E]) at positions 7 and 8.

In summary, the data described in this section define, by the use of a quantitative assay, the structural requirements for the A2.1 anchor positions 2 and 9. Furthermore, they indicate that several other positions also have significant impact in determining the A2.1 binding capacity of a peptide ligand.

Anchor Residues Are Not Sufficient to Determine A2.1 Binding

To probe further the role of nonanchor positions in A2.1 binding, a selection of 18 peptides of various viral origins, all sharing the L_2V_9 motif, were synthesized and tested for A2.1 binding (Table 1). It was found that binding affinity, despite the presence of identical main anchor residues, varied in this set of naturally occurring peptide sequences over a >10,000-fold range, from 7 nM to greater than 50 μM. To expand this analysis, a larger library of motif-containing peptides was generated. The sequences of various viral and tumor antigens were scanned for the presence of 9-mer and 10-mer peptides containing canonical 2/9 or 2/10 A2.1 motifs, as defined above. By this approach, a total of 161 9-mer peptides and 170 10-mer peptides containing an L or M in position 2 and a V, I, or L in the C-terminal (position 9 or 10) were selected, synthesized, and tested for binding to HLA-A2.1 molecules.

In good agreement with the data shown in Table 1, large variations in A2.1 binding affinity were detected. More specifically, as shown in Table 2, only 11.8% of the 9-mer peptides were high affinity binders (K_D, <50 nM), and 22.4% of the peptides bound with intermediate affinity (K_D, 50–500 nM). Of the peptides in this set, 36% were weak binders (K_D, 0.5–50 μM), and 29.8% had no detectable binding (K_D, >50 μM). It should be noted that the choice of the different binding categories used in this study were derived from two sets of experimental observations. First, ~80% of the known HLA-restricted epitopes examined so far bind in our "good" category, with a K_D of ≤50 nM, and ~20% in the 50–500 nM range (A. S. et al., submitted). Second, ~90% of MHC-bound naturally processed peptides tested so far bind with a K_D of ≤50 nM, with ~10% in the 50–500 nM range (A. S. and V. H. Engelhard, unpublished data). Thus, the data to date would indicate that good and (to a lower extent) intermediate binding are associated with functional activity.

In the 10-mer set of peptides, we found a lower percentage of peptides binding with high (5.9%) and intermediate

Table 1. A2.1 Binding of Peptides of Viral or Tumor Origin Containing Canonical 2/9 Anchor Residues

Source	Peptide Sequence									50% Inhibitory Dose (nM)
HBV	W	L	S	L	L	V	P	F	V	6.9
HCV	Y	L	V	A	Y	Q	A	T	V	45
HBV	S	L	Y	A	V	S	P	S	V	64
c-erb2	H	L	Y	Q	G	C	Q	V	V	147
HIV	L	L	W	K	G	E	G	A	V	217
p53	L	L	G	R	N	S	F	E	V	357
HCV	Y	L	V	T	R	H	A	D	V	454
HIV	I	L	K	E	P	V	H	G	V	909
HCV	G	L	R	D	L	A	V	A	V	1,470
c-erb2	P	L	T	S	I	I	S	A	V	3,333
HIV	H	L	E	G	K	V	I	L	V	8,333
HBV	H	L	S	L	R	G	L	P	V	12,500
HPV	R	L	C	V	Q	S	T	H	V	16,666
HPV	E	L	R	H	Y	S	D	S	V	–
c-erb2	D	L	A	A	R	N	V	L	V	–
c-erb2	V	L	V	K	S	P	N	H	V	–
HIV	E	L	H	P	D	K	W	T	V	–
HBV	C	L	T	F	G	R	E	T	V	–

As described in Experimental Procedures, 9-mer peptides containing the L_2V_9 anchors were selected from viral or tumor protein sequences and tested for A2.1 binding. The binding is expressed as the nanomolar dose of peptide that yielded 50% inhibition of the standard peptide. The average of at least two independent experiments is shown. A dash indicates an inhibitory dose of >50,000 nM.

Abbreviations: HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus.

(17.1%) affinity. This set of peptides also contained 41.2% weak binders, and 35.9% were nonbinders. Taken together, these results indicate that 10-mer peptides also bind to A2.1, if with a somewhat lower average affinity. Most importantly, the high number of weak and nonbinding peptides in both sets of peptides (approximately 65%–75% of the total) indicate that factors other than size and anchor residues are also critical in determining whether or not a peptide will bind A2.1 MHC molecules.

The Role of Nonanchor Residues in Determining A2.1 Binding

To define more specifically what factors other than size and anchor residues are important for A2.1 binding capacity, the following analysis was performed. We reasoned that amino acid residues that facilitated high affinity binding would be overrepresented in high affinity peptides, whereas amino acid residues that were detrimental to high affinity binding would be underrepresented in high affinity binders and overrepresented in nonbinders. Accordingly, we calculated the frequency of each amino acid in each of the nonanchor positions for the 55 good and intermediate binders and the 48 nonbinders found in the 9-mer peptide set. A similar analysis was also performed for the 39 good and intermediate binders versus the 61 nonbinders found in the 10-mer peptide set. Amino acids of similar characteristics were grouped together to increase the number of data points in each category. An example of this analysis for position 3 in the 9-mer peptide set is shown in Table 3. Several striking trends are revealed. For example, the frequency of aromatic residues (tyrosine [Y], F, and tryptophan [W]) was greatly increased in binding, as compared with nonbinding peptides (21.8% versus 4.2%). Conversely, while 12.5% (six peptides) of the nonbinders carried a negative charge (residues D or E) in position 3, these same residues were completely absent in this position in the subset of high and intermediate A2.1-binding peptides. Similarly, positively charged residues (R, K, and histidine [H]) were much more frequent in nonbinders than in binders (16.7% versus 3.6%).

To quantitate these differences more conveniently, the frequency of a given amino acid group in A2.1 binders was divided by the frequency in nonbinders to obtain a frequency ratio. This ratio indicates whether a given group of residues occurs at a given position preferentially in binders (ratio, >1) or in nonbinders (ratio, <1). Frequency ratios are shown for position 3 in Table 3. A complete listing for all nonanchor positions in 9-mer and 10-mer peptides is shown in Figure 2.

To facilitate analysis of the data shown in Figure 2, a

Table 2. A2.1 Binding Capacity of Motif Containing 9-mer and 10-mer Peptides

Binding Capacity (K_d)	Number of Peptides (%)	
	9-mers	10-mers
Good (≤50 nM)	19 (11.8%)	20 (5.9%)
Intermediate (50–500 nM)	36 (22.4%)	29 (17.1%)
Weak (500 nM–50 μM)	58 (36%)	70 (41.2%)
Negative (>50 μM)	48 (29.8%)	61 (35.9%)
Total	161 (100%)	170 (100%)

Peptides containing the canonical anchors L or M in position 2 and V, L, or I in the C-terminus (position 9 or 10) were selected from protein sequences of viral or tumor origin, synthesized, and tested for A2.1 binding, and the 50% inhibitory concentration was calculated. Peptides were separated according to their binding capacity into the indicated groups. The numbers of peptides belonging to each group and their percentage of the total number of peptides tested are given.

Table 3. Frequency of Various Amino Acid Groups in Position 3 of 9-mer Peptides

Amino Acid Group	Good and Intermediate Binders: Number of Peptides (%)	Nonbinders: Number of Peptides (%)	Ratio
Y, F, W	12 (21.8)	2 (4.2)	5.2
S, T, C	11 (20)	5 (10.4)	1.9
L, V, I, M	19 (34.5)	9 (18.8)	1.8
A	2 (3.6)	2 (3.6)	0.9
P	4 (7.3)	5 (10.4)	0.7
G	2 (3.6)	4 (8.3)	0.4
Q, N	3 (5.5)	7 (14.6)	0.4
R, H, K	2 (3.6)	8 (16.7)	0.2
D, E	0 (0)	6 (12.5)	0.0
Total	55 (100)	48 (100)	

An example is given for the calculation of ratios to determine association of certain residues with good or poor binding. One hundred sixty-one 9-mer peptides containing the canonical anchors L or M in position 2 and V, L, or I in position 9 were selected from protein sequences of viral or tumor origin, synthesized, and tested for A2.1 binding. From this set, the 58 good and intermediate binders, as well as the 48 peptides with no detectable binding, were compared in this analysis. The number of peptides in these two groups that contained a certain residue in position 3 and the percentage of occurrence of a certain residue are given. From the values obtained for the two groups, a ratio was calculated for each residue or group of residues to indicate whether a residue occurs preferentially in good/intermediate binders (ratio, >1) or nonbinders (ratio, <1).

threshold level was set for the ratios, such that residues that had a more than 4-fold greater frequency in binders compared with nonbinders were regarded as favored or preferred residues and residues that had a more than 4-fold lower frequency in binders than in nonbinders were regarded as unfavored or deleterious residues. Following this approach, groups of residues showing prominent associations with either A2.1 binding capacity or lack thereof were identified (Figure 3). In general, the most detrimental effects were observed with charged amino acids. At position 1, both P and acidic (E and D) residues were infrequent in A2.1-binding peptides. At position 6, basic (H, R, and K) residues were associated with nonbinding peptides, whereas both acidic and basic residues were infrequent in good binders at positions 3 and 7. Conversely, aromatic residues were associated with high affinity binding in positions 1, 3, and 5. Furthermore, residues with OH- or SH-containing side chains, such as S, T, or C, were favored at position 4, while A was favored in position 7 and P in position 8. In conclusion, frequency analysis of different amino acid groups allowed us to define a more accurate A2.1 motif that takes into account the impact of positions other than anchor positions 2 and 9 on A2.1 binding affinity.

Analysis of 10-mer A2.1 Ligands

The same approach described above for 9-mer peptides was also used to analyze the data obtained with the set of 10-mer peptides (Figures 2B and 3). At the N- and C-termini of the peptides, the pattern observed was rather simi-

A 9-mer Peptides

	1	2	3	4	5	6	7	8	9	
YF,W	7.0		5.2	0.9	8.7	2.0	2.3	2.6		
S,T,C	0.7	A	1.9	4.8	0.9	1.2	1.2	1.1	A	
L,V,I,M	3.1	N	1.8	0.5	0.9	1.3	1.2	1.7	N	
A	2.6	C	0.9	0.9	0.7	0.9	4.4	0.3	C	
P	0.1	H	0.7	0.7	2.6	1.7	2.9	∞	H	
G	3.5	O	0.4	1.1	1.1	1.3	0.4	0.4	O	
Q,N	0.5	R	0.4	1.2	0.9	1.0	0.7	0.3	R	
R,H,K	3.1		0.2	1.0	0.9	0.1	0.0	1.3		
D,E	0.1		0.0	0.7	0.3	0.7	0.1	0.9		

B 10-mer Peptides

	1	2	3	4	5	6	7	8	9	10	
L,V,I,M	3.0		10.2	1.0	1.3	2.1	1.4	4.7	0.8		
A	∞	A	3.1	0.2	1.6	0.6	1.3	1.8	0.9	A	
Y,F,W	∞	N	2.8	2.1	3.6	0.8	1.8	14.1	2.1	N	
Q,N	1.0	C	0.9	0.8	0.8	0.8	0.8	0.4	0.7	C	
S,T,C	0.9	H	0.8	1.1	1.0	0.9	1.4	1.3	2.8	H	
G	0.8	O	0.5	4.7	0.8	6.3	2.7	0.7	0.8	O	
P	0.0	R	0.4	2.8	0.0	1.0	0.4	1.9	1.2	R	
R,H,K	1.2		0.3	0.1	0.7	0.4	0.2	0.0	0.2		
D,E	0.0		0.2	0.6	0.3	1.0	0.3	0.0	0.4		

Figure 2. Relative Frequency of Amino Acid Groups at Different Nonanchor Positions in 9-mer and 10-mer Peptides

Peptides containing the canonical anchors L or M in position 2 and V, L, or I in the C-terminus (position 9 or 10) were selected from protein sequences of viral or tumor origin, synthesized, and tested for A2.1 binding. The frequency of occurrence of certain amino acids or groups of amino acids was determined for each position, and frequency ratios in good/intermediate binders or nonbinders were calculated, as exemplified in Table 3. Ratios that differed from unity by greater than 4-fold are bolded.

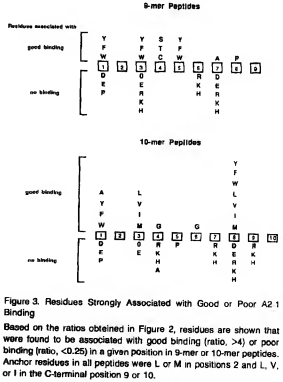

Figure 3. Residues Strongly Associated with Good or Poor A2.1 Binding

Based on the ratios obtained in Figure 2, residues are shown that were found to be associated with good binding (ratio, >4) or poor binding (ratio, <0.25) in a given position in 9-mer or 10-mer peptides. Anchor residues in all peptides were L or M in positions 2 and L, V, or I in the C-terminal position 9 or 10.

Table 4. Validation of the Extended Motifs Using an Independent Set of 9-mer and 10-mer Peptides

Peptides	Source	Position	Sequence										50% dose (nM)
Peptides Carrying Favorable Residues													
9-mers	HBV	1406	S	L	**Y**	A	D	S	P	S	V		14
	HBV	204	**F**	L	G	G	T	T	V	C	L		29
	HBV	1266	F	L	C	K	**Q**	Y	L	N	L		29
	HBV	1045	G	L	S	R	**Y**	V	A	R	L		42
	HIV	280	R	M	**Y**	**S**	P	I	S	I	L		56
	MAGE	101	A	L	S	R	K	V	**A**	E	L		100
10-mers	HBV	437	**F**	L	P	S	D	F	F	P	S	V	3.3
	HBV	484	L	M	T	L	A	T	W	**V**	G	V	7.4
	HBV	408	**F**	L	P	S	D	F	F	P	S	I	9.4
	HBV	478	I	L	C	W	G	E	L	**M**	T	L	26
	HIV	2181	N	L	**W**	V	T	V	Y	**Y**	G	V	31
	HIV	2181	K	L	**W**	V	T	V	Y	**Y**	G	V	33
	HBV	1306	**R**	M	**R**	**G**	T	F	S	**A**	**P**	L	45
	MAGE	150	S	L	Q	L	V	F	G	**I**	E	L	102
	HBV	246	F	L	C	I	L	L	L	C	L	I	2,380
Peptides Carrying Deleterious Residues													
9-mers	HBV	1154	H	L	E	**S**	L	F	T	A	V		10,000
	HBV	1052	L	L	**S**	**S**	**N**	S	R	I	L		10,000
	HBV	1420	D	L	V	H	**F**	A	S	P	L		16,667
	PAP	201	D	L	F	G	I	**W**	S	K	V		25,000
	MAGE	147	S	L	R	V	L	A	A	G	L		–
	HBV	511	N	M	G	L	K	F	R	Q	L		–
	HBV	83	D	L	V	V	S	Y	V	N	V		–
	MAGE	103	D	L	Q	Q	S	L	R	V	L		–
	PAP	352	P	L	E	R	F	A	E	L	V		–
10-mers	MAGE	93	D	L	E	**S**	E	F	Q	A	**A**	L	–
	MAGE	101	A	L	**S**	R	K	V	A	E	L	V	–
	HBV	511	N	M	G	L	K	F	R	Q	L	L	–
	HBV	805	P	L	S	Y	Q	H	F	R	R	L	–
	HBV	1069	N	L	H	D	S	C	S	R	N	L	–
	HBV	1084	L	L	Y	Q	T	F	G	R	K	L	–
	HIV	2181	D	L	**W**	V	T	V	Y	Y	G	V	–
	MAGE	135	K	L	P	G	L	L	S	R	D	L	–

Two independent sets of 9-mer and 10-mer peptides not previously analyzed were studied for binding. The 50% inhibitory dose of these peptides represents the average of at least two independent experiments. A dash indicates a 50% dose of greater than 50,000 nM. Peptides were sorted by whether they contained a deleterious residue or not. Residues associated with poor binding, as listed in Figure 3, are underlined, and residues associated with good binding are shown in bold.

Abbreviations: HBV, hepatitis B virus; HIV, human immunodeficiency virus; MAGE, melanoma-associated gene expression antigen; PAP, prostatic acid phosphatase.

lar to the one observed with 9-mers. For instance, in the 10-mer set, as in the case of the 9-mer peptides, position 1 was characterized by an increased frequency of aromatic residues in the binder set, while negative charges and P were again associated with poor binding. Again at position 3, negative charge was associated with poor binding. Interestingly, at this position, aliphatic (rather than aromatic) residues were associated with high affinity binding. At the C-termini of the peptides, certain similarities were also observed. In the 10-mer, the penultimate residue at position 9 (corresponding to position 8 in the 9-mer) was quite permissive, with only basic residues being found more frequently in nonbinders. Similar to the situation at position 7 in the 9-mer, neither positive nor negative charges were tolerated in the antepenultimate position 8 of the 10-mers. Also, position 7 did not favor positive residues in the 10-mers, as previously observed for position 6 in the 9-mers. In analogy to what was observed at position 3, the residues associated with good binding were, however, different. Aromatic and hydrophobic residues were frequent in high affinity binders at position 8 (as opposed to only A being frequent at position 7 in the 9-mers).

Finally, a rather distinctive pattern was observed in the middle of the peptide. At position 4, G was favored in high binders, while both A and positive charges were very frequent in nonbinders. P, in position 5, was completely absent in the A2.1 binders. It is noteworthy that none of the trends observed in positions 4 and 5 in the 10-mer set have any counterpart in position 4 or 5 in the 9-mer set.

In conclusion, a detailed motif can be generated for A2.1 10-mer peptides, following a strategy similar to the one described for 9-mer peptides above. Both important differences and striking similarities can be noted in comparing the 9-mer and 10-mer sets at these nonanchor positions.

Validation of A2.1 Binding Requirements Using an Independent Peptide Set

The value of the extended motifs defined above for pre-

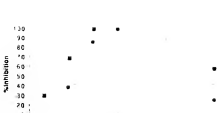

Figure 4. A2.1 Binding of Consensus Peptides and Poly(A) Peptides

Nonamer and decamer consensus peptides and poly(A) analogs, both containing the same canonical L_2V_9 anchors, were tested in a peptide dose titration for the inhibition of the A2.1 binding of the standard peptide.

dicting HLA-A2.1 binding was examined using independent panels of 9-mer and 10-mer peptides (Table 4). For this purpose, the following two sets were synthesized: 15 peptides (of either 9 or 10 residues in length) that contained no unfavored residues and at least one favored residue, and a set of 17 peptides (9-mers or 10-mers) that carried at least one unfavored residue in their sequence. All peptides contained canonical anchor residues in position 2 (L or M) and position 9 (L, V, or I). When their binding capacity was determined, it was found that 11 of 15 peptides carrying no unfavored residues and one or more favored residues bound with high affinity (≤50 nM). Three bound in the 50–100 nM range, and only one bound in the micromolar range. In contrast, none of the peptides carrying an unfavored residue bound HLA-A2.1 with high affinity, with only four peptides demonstrating any detectable binding in the 10–25 μM range. Interestingly, peptides carrying both one favored residue and one or more unfavored residues were found to be poor or negative binders.

Generation of High Affinity A2.1 Consensus Binders

The data presented in the previous sections demonstrate that, besides the two main anchors located at position 2 and at the C-terminus of a peptide, the residues present in most other positions influence, either in a positive or negative way, the A2.1 binding affinities of peptides. As previously mentioned, the poly(A)-based A2.1 model 9-mer peptide (ALAKAAAAV) bound with intermediate affinity in the 50–500 nM range (50% inhibitory concentration, 170 nM ± 72). An analogous 10-mer peptide (ALAKAAAAAV) also bound A2.1, but with approximately 10-fold lower affinity (50% inhibitory concentration, 1667 nM ± 778). It was considered likely that these poly(A) peptides bind because they have the appropriate main anchors, but that their affinity is relatively low because they lack favored residues. If this reasoning was correct, then insertion of favored amino acids at each of the nonanchor positions should result in peptides that bind much better than the simple poly(A) peptides. To test this prediction, two such peptides, one of 9 and one of 10 residues, were synthesized, and their A2.1 binding capacity was measured. Representative data are shown in Figure 4. Both of these preferred consensus peptides bound with high affinities, with a 50% inhibitory concentration of 1.3 and 6.2 nM. These values correspond to more than a 130-fold increase in affinity for the 9-mer and a 270-fold increase in affinity for the 10-mer, as compared with the simple poly(A) peptides.

Overall Correlation between Presence of Expanded A2.1 Motifs and Binding Capacity

A compilation of all the data available to date for analysis is shown in Table 5. A total of 422 peptides, all carrying canonical anchor residues in positions 2 and 9, have been tested for A2.1 binding. Peptides were divided into three groups: no detrimental residues and one or more favored residues, no detrimental or favored residues, and one or more detrimental residues. The frequencies of good and intermediate binders (K_D, ≤500 nM), weak binders (K_D, 0.5–50 μM), and nonbinders in each peptide group are shown in Table 5. Of the 121 peptides that contained no unfavored and at least one favored residue(s), 83 (69%) were good or intermediate binders. By contrast, of the 227 peptides carrying one or more unfavored residues, only 19 (8%) bound A2.1 with high or intermediate affinity. The third, smaller group of peptides that carry neither favored nor unfavored residues was found to contain approximately equal numbers of good and intermediate, weak, and negative binders. Thus, the value of this expanded motif for predicting high and intermediate binding peptides of 69% is more than twice that of the 30.3% (128 of 422) found when these peptides were analyzed only on the basis of their containing the canonical 2/9 anchor motif.

Discussion

In the present report, an extensive analysis of the interaction between A2.1 molecules and their peptide ligands has been presented. This analysis, rendered possible by the development of high throughput quantitative binding assays (A. S. et al., submitted) utilizing synthetic peptides and purified A2.1 molecules, allows for the definition of the functional roles of the main anchor residues, as well as other positions (secondary anchor residues) of peptide ligands in the interaction between peptides and A2.1 molecules.

The Role of Positions 2 and 9 as Anchor Residues in A2.1 Binding

In good agreement with previous reports by several independent groups (Falk et al., 1991; Hunt et al., 1992; Parker et al., 1992), it was found that two aliphatic hydrophobic anchor residues located at position 2 and at the C-terminus of peptides 9 or 10 residues in length were essential for A2.1 binding. The availability of quantitative binding assays allowed an analysis of the importance of the two main anchors relative to each other and quantitation of the impact on binding affinity of substituting these two anchor residues with a diverse set of side chain replacements. Using a model poly(A)-based A2.1-binding peptide, it was found that the two main anchors have similar degrees of stringencies. At position 2 only L and M and at the C-terminus only V, I, and L were associated with optimal binding capacity, and only very conservative substitutions,

Table 5. Correlation of A2.1 Binding and the Presence of Preferred or Deleterious Residues

Residues in Nonanchor Positions					
Preferred	Detrimental	Binders (K_d, ≤500 nM)	Weak Binders (K_d, ≤50 μM)	Nonbinders (K_d, >50 μM)	Total
≥1	0	83	33	5	121
0	0	25	32	16	74
0 or ≥1	≥1	19	82	126	227
Total		128	147	147	422

All peptides tested in this study containing canonical anchors in position 2 and the C-terminus were grouped according to the presence of preferred residues, deleterious residues, or both, as defined in Figure 3, and according to their binding capacity, as indicated.

such as I, V, A, and T at position 2 (or A and M in position 9), were still compatible with binding, albeit with a 10- to 100-fold lower affinity. All other residues tested effectively obliterated the binding capacity of the peptide. Thus, these experiments confirmed the crucial role of anchor positions 2 and 9 in A2.1 binding and more exactly defined the specificity of the binding to the A2.1 molecule. However, when large numbers of naturally occurring peptides, all carrying optimal anchor residues in positions 2 and 9 (or 10), were synthesized and tested for binding, it was found that the measurable affinities varied over a 10,000-fold range (5 nM to 50 μM), with approximately a third of the peptides tested not showing any detectable A2.1 binding capacity. These findings extend previous observations by others (Parker et al., 1992; Jameson and Bevan, 1992), as well as confirm data obtained with single amino acid substitutions of the poly(A) model A2.1 binder in positions other than 2 and 9 (Figure 2) that strongly suggested that factors other than peptide size and the presence of the two main anchor residues play a decisive role in determining A2.1 binding capacity.

Definition of Expanded A2.1 Motifs (A2.1 Secondary Anchor Residues)

By analyzing the correlation between the sequences of large numbers of peptides containing the primary anchor residues and their A2.1 binding capacity, prominent secondary effects have been defined. Two different lines of reasoning can be invoked to explain these effects. First, it is likely that some of the residues lead to negative effects, either conformational effects, steric hindrances, or repulsive electrostetic interactions. In fact, it is interesting to note that charged residues were responsible for most of the deleterious effects, while most of the positive associations with A2.1 binding were detected with hydrophobic or eromatic residues. A simlier dominence of negetive effects of charged amino acid substitutions on peptide binding has recently been implicated by Boehncke et el. (1993) in a murine class II system and by our own group in the DR4w4 system (Sette et al., 1993).

Most of the effects detected in position 1 of the peptide are easily interpreteble along these lines, in the context of the known X-ray crystallogrephic structure of A2.1 molecules (Saper et al., 1991) and of peptide–class I complexes in general (Travers and Thorpe, 1992). More specifically, pocket A at the end of the peptide binding groove is known to engege the positively cherged N-terminus of the peptide through hydrogen bonding of several conserved Y residues (Latron et al., 1992). In this study we have found that neither negative residues nor P is allowed in position 1. Introduction of negatively charged residues in positon 1 might disrupt the electrostatic environment of this pocket, while the backbone distortion induced by the presence of a P in position 1 might interfere with the required positioning of the NH_3^+ group necessary for good hydrogen bonding.

The second mechanism that might explain the prominent secondary effects, especially the ones observed in positions 3, 6, and 7 (7 and 8 in the 10-mers), is that these residues may act as secondary anchors for A2.1 binding. In retrospect, the existence of secondary anchor residues in A2.1 binding could have been predicted from the crystallogrephic studies based on the existence of several secondary pockets within the A2.1 molecule that are occupied by extra electron density, presumably derived from side chains of self-peptides (Saper et al., 1991). Specifically, based on the published data, it is likely that the D pocket might be engaged in interactions with some of the residues present in position 3 of the peptide, while residues 6 and 7 (7 or 8 in the 10-mers) might be involved in interactions with the A2.1 region corresponding to the C and E pockets.

Interestingly, the secondary effects revealed by our analysis differ in the case of 9-mer versus 10-mer peptides, especially in the middle section of the peptide. This finding is in good agreement with crystallographic analyses that have shown that 9- and 10-mer peptides binding to the same class I molecule have similar main anchor points, but "bulge" differently in the middle (Guo et al., 1992; Parham, 1992). In this report, it is also interesting to note that positions 4 and 5 in the 9-mers (5 end 6 in the 10-mers) are also among the most permissive ones. These residues, because of their bulging out away from the class I binding cleft end because of their permissiveness, are thus ideally suited to function as prominent T cell receptor contact sites. Because these secondary effects differ between 9-mers and 10-mers, in retrospect it is easy to explain why definition of secondery anchors by pool sequencing has not been possible.

Finelly, we would like to comment on the practical implications of the date presented herein. Our data show that definition of accurate binding motifs, based on both primary end secondary anchors, allows a high predictability of class I binding. Since antigenicity and binding are closely correleted (Buus et el., 1987; Schaeffer et al.,

1989), an accurate motif also allows a better prediction of effective immunogens and promises to constitute a useful tool in the development of peptide-based vaccines.

Experimental Procedures

A2.1 Purification

The HLA-A2.1-positive human Epstein–Barr virus–transformed B cell line JY was used as a source of A2.1 molecules, purified as described elsewhere (A. S. et al., submitted). In brief, cell lysates from large-scale (10^9–10^{11}) cell cultures were filtered through 0.45 μM filters and purified by affinity chromatography. Columns of inactivated Sepharose CL4B and protein A–Sepharose were used as precolumns. The cell lysate was subsequently depleted of HLA-B and HLA-C molecules by repeated passege over protein A–Sepharose beads conjugated with the anti-HLA-B/C antibody B1.23.2 to remove HLA-B and HLA-C molecules. Subsequently, the anti-HLA-A/B/C antibody W6/32 was used to capture HLA-A2.1 molecules. Molecules were eluted with diethylamine, 1% n-octyl glucoside (pH 11.5), neutralized with 1 mM Tris (pH 6.8), concentrated by ultrafiltration on Amicon 30 cartridges, and stored at 4°C. Protein purity, concentration, and effectiveness of depletion steps were monitored by SDS–polyacrylemide gel electrophoresis.

Peptide Synthesis

Peptides were synthesized either at Cytel Corporetion or, for large epitope libraries, by Chiron Mimotopes (Chiron Corporation). Peptides prepered at Cytel were synthesized by sequential coupling of N-α-Fmoc-protected amino acids on an Applied Biosystems 430A peptide synthesizer using standard Fmoc-coupling cycles (software version 1.40). All amino acids, reagents, and resins were obteined from Applied Biosystems or Bachem. Solvents were obtained from Burdick & Jackson. Solid-phase synthesis wes started from an appropriately substituted Fmoc–amino ecid–Sasrin resin. The loading of the sterting resin was 0.5–0.7 mmol/g polystyrene, and 0.1 or 0.25 meq were used in each synthesis. A typical reaction cycle proceded as follows. The N-terminal Fmoc group was removed with 25% piperidine in dimethylformemide for 5 min, followed by another treatment with 25% piperidine in dimethylformamide for 15 min. The resin wes washed five times with dimethylformamide. An N-methylpyrolidone solution of a 4- to 10-fold excess of e preformed 1-hydroxybenzotriazole ester of the appropriate Fmoc–amino acid was added to the resin, and the mixture was allowed to react for 30–90 min. The resin was weshed with dimethylformemide in preparation for the next elongation cycle. The fully protected, resin-bound peptide was subjected to e piperidine cycle to remove the terminal Fmoc group. The product was washed with dichloromethane and dried. The resin was then treated with trifluoroacetic acid in the presence ol appropriete scevengers (e.g., 5% [v/v] in water) for 60 min at 20°C. After evaporation of excess trifluoroacetic acid, the crude peptide was washed with diethyl ether, dissolved in water, and lyophilized. The peptides were purified to >95% homogeneity by reverse-phase high pressure liquid chromatography using H_2O/CH_3CN gradients conteining 0.2% trifluoroacetic acid modifier on a Vydac 300 Å pore-sized C-18 preparative column. The purity of the synthetic peptides was assayed on an analytical reverse-phase column and their composition ascertained by amino ecid anelysis, sequencing, or both.

A2.1 Binding Assay

A quantitative assay for A2.1-binding peptidas besed on the inhibition of binding of a radiolebeled stenderd peptide to detergent-solubilized MHC molecules is described elsewhere (A. S. et al., submitted). In brief, tha HBc 18–27 F→Y6 (FLPSDYFPSV) peptide wes iodinated by the chloramine T method (Buus at al., 1987). MHC concentrations yielding approximately 15% of bound peptide were used in the inhibition assays (usually in the 10 nM renge). Various doses of the test peptides (ranging from 100 μM to 1 nM) were coincubated, together with the ~ 5 nM radiolabelad HBc 18–27 peptida and A2.1 molecules, for 2 days at room temparature in the presence of a cocktail of protease inhibitors and 1 μM β$_2$-microglobulin. The final concentrations of peptide inhibitors were 1 mM phenylmethylsulfonyl fluoride, 1.3 mM 1,10 phenenthroline, 73 μM pepstatin A, 8 mM EDTA, and 200 μM N-α-p-tosyl-L-lysine chloromethyl ketone. Final detergent concentration in the mixture was 0.05% Nonidet P-40. At the end of the incubation period, the percent of MHC-bound radioactivity wes determined by gel filtration, and the 50% inhibitory dose was calculeted for each peptide, as described (Sette et al., 1992).

Acknowledgments

Correspondence should be addressed to A. S. We thank Merie-France del Guercio and Scott Southwood for their technical assistence and Joyce Joseph for her assistence in preparation of the manuscript.

Received May 26, 1993; revised June 22, 1993.

References

Bjorkman, P. J., and Parham, P. (1990). Structure, function, and diversity of class I major histocompatibility complex molecules. Annu. Rev. Biochem. *59*, 253–288.

Boehncke, W.-H., Takeshita, T., Pendleton, C. D., Houghten, R. A., Sadegh-Nasseri, S., Racioppi, L., Berzofsky, J. A., and Germein, R. N. (1993). The importance of dominent negative effects of amino acid side chain substitution in peptide–MHC molecule interactions and T cell recognition. J. Immunol. *150*, 331–341.

Buus, S., Sette, A., Colón, S., Miles, C., and Grey, H. M. (1987). The relation between major histocompatibility complex (MHC) restriction and the capacity of le to bind immunogenic peptides. Science *235*, 1353–1358.

Chen. B. P., and Perham, P. (1989). Direct binding of influenze peptides to class I HLA molecules. Nature *337*, 743–745.

Corr, M., Boyd, L. F., Frenkel, S. R., Kozlowski, S., Peclan, E. A., and Margulies, D. H. (1992). Endogenous peptides of a soluble major histocompatibility complex class I molecule, H-2Ld: sequence motif, quantitative binding, and molecular modeling of complex. J. Exp. Med. *176*, 1681–1692.

Elliot, T., Cerundolo, J. E., and Townsend, A. (1991). Peptide-induced conformational change of the class I heevy chein. Neture *351*, 402–406.

Falk, K., Rötzschke, O., Stefanovic, S., Jung, G., and Rammensee, H.-G. (1991). Allele specific motifs revealed by sequencing of self peptides eluted from MHC molecules. Neture *351*, 290–296.

Fremont, D. H., Matsumura, M., Sture, E. A., Peterson, P. A., and Wilson, I. A. (1992). Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. Science *257*, 919–927

Garret, T. P. J., Bjorkman, P. J., and Wiley, D. C. (1989). Specificity pockets for the side chains of peptide antigens in HLA-Aw68. Neture *342*, 692–696.

Guo, H. C., Jardetzky, T. S., Garrat, T. P. J., Lane, W. S., Strominger, J. L., and Wiley, D. C. (1992). Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. Nature *360*, 364–366.

Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., and Engelhard, V. H. (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science *255*, 1261–1266.

Jameson, S. C., and Bevan, M. J. (1992). Dissection of major histocompatibility complex (MHC) end T cell receptor contact residues in e Kb-restricted ovalbumin peptida and an assessment of the predictive power of MHC-binding motifs. Eur. J. Immunol. *22*, 2663–2667.

Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R., and Wiley, D. C. (1991). Identification of self peptides bound to purified HLA-A27. Nature *353*, 326–329.

Latron, F., Pazmany, L., Morrison, J., Moots, R., Saper, M. A., McMichael, A., and Strominger, J. L. (1992). A critical role for conserved residues in the cleft of HLA-A2 in presentation of e nonepeptide to T cells. Science *257*, 964–967.

Matsumura, M., Fremont, D. H., Peterson, P. A., and Wilson, I. A. (1992). Emerging principles for the recognition of peptide antigens by MHC class I molecules. Science *257*, 927–934.

Parhem, P. (1992). Deconatructing the MHC. Nature *360*, 300–301

Parker, K. C., Bednarke, M. A., Hull, L. K., Utz, U., Cunningham, B., Zweerink, H. J., Biddison, W. E., and Coligan, J. E. (1992). Sequence motifs important for peptide binding to the human MHC class I molecule, HLA-A2. J. Immunol. *149*, 3580–3587.

Rötzschke, O., and Falk, K. (1991). Naturally-occurring peptide antigens derived from the MHC class I–restricted processing pathway. Immunol. Today *12*, 447–455.

Rötzschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, G., and Rammensee, H. G. (1990). Isolation and anelysis of naturally processed peptides as recognized by cytotoxic T cells. Nature *348*, 252–254.

Saper, M. A., Bjorkman, P. J., and Wiley, D. C. (1991). Refined structure of the human histocompatibility antigen HLA-A2 at 26 Å resolution. J. Mol. Biol. *219*, 277–319.

Schaeffer, E. B., Sette, A., Johnson, D. L., Bekoff, M. C., Smith, J. A., Grey, H. M., and Buus, S. (1989). Relative contribution of "determinant selection" and "holes in the T cell repertoire" to T cell responses. Proc. Natl. Acad. Sci. USA *86*, 4649–4653.

Schumacher, T. N. M., De Bruijn, M. L. H., Vernie, L. N., Kast, W. M., Melief, C. J. M., Neefjes, J. J., and Ploegh, H. L. (1991). Peptide selection by MHC class I molecules. Nature *350*, 703–706.

Sette, A., Southwood, S., O'Sullivan, D., Gaeta, F. C. A., Sidney, J., and Grey, H. M. (1992). Effect of pH on class II–peptide interactions. J. Immunol. *148*, 844–851.

Sette, A., Sidney, J., Oseroff, C., del Guercio, M.-F., Southwood, S., Arrhenius, T., Powell, M. F., Colón, S. M., Gaeta, F. C. A., and Grey, H. M. (1993). HLA DR4w4 binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide–DR interactions. J. Immunol., in press.

Townsend, A., and Bodmer, H. (1989). Antigen recognition by class I–restricted T lymphocytes. Annu. Rev. Immunol. *7*, 601–624.

Travers, P. J., and Thorpe, C. J. (1992). Anchored fore and aft. Curr. Biol. *2*, 679–681.